

Using high-frequency SCADA data for wind turbine performance monitoring: a sensitivity study

Elena Gonzalez^{a,b}, Bruce Stephen^b, David Infield^b, Julio J. Melero^{a,*}

^a*CIRCE-Universidad de Zaragoza, C/Mariano Esquillor Gómez 15, 50018 Zaragoza, Spain*

^b*Department of Electronic and Electrical Engineering, University of Strathclyde, 204 George Street, Glasgow G1 1XW, UK*

Abstract

Intensive condition monitoring of wind generation plant through analysis of routinely collected SCADA data is seen as a viable means of forestalling costly plant failure and optimising maintenance through identification of failure at the earliest possible stage. The challenge to operators is in identifying the signatures of failure within data streams and disambiguating these from other operational factors. The well understood power curve representation of turbine performance offers an intuitive and quantitative means of identifying abnormal operation, but only if noise and artefacts of operating regime change can be excluded. In this paper, a methodology for wind turbine performance monitoring based on the use of high-frequency SCADA data is employed featuring state-of-the-art multivariate non-parametric methods for power curve modelling. The model selection considerations for these are examined together with their sensitivity to several factors, including site specific conditions, seasonality effects, input relevance and data sampling rate. The results, based on operational data from four wind farms, are discussed in a practical context with the use of high frequency data demonstrated to be beneficial for performance monitoring purposes whereas further attention is required in the area of expressing model uncertainty.

Keywords: Wind turbine, Power curve, High-frequency data, Performance

*Corresponding author

Email addresses: elena.gongar@gmail.com (Elena Gonzalez), melero@unizar.es (Julio J. Melero)

monitoring, SCADA data

List of abbreviations

ANN Artificial Neural Network

CM Condition Monitoring

k -NN k -Nearest Neighbours

5 **MAE** Mean Absolute Error

MOB Method of bins

O&M Operation and Maintenance

OOB out-of-bag

PC Power Curve

10 **RF** Random Forest

RMSE Root Mean Square Error

SCADA Supervisory Control And Data Acquisition

SVR Support Vector Regression

WF Wind Farm

15 **WT** Wind Turbine

1. Introduction

Wind energy is expected to become the largest source of renewable electricity by 2020 [1]. However, the sector faces challenges relating to the Operation and Maintenance (O&M) of existing installations. The lack of certainty on revenue
20 stability has put pressure on reducing O&M costs, that can actually reach up to

30% in offshore environments [2]. There is therefore an ever-increasing interest in optimising the O&M strategies in both the industry and the academia.

Condition Monitoring (CM) systems are increasingly installed with the goal of providing Wind Turbine (WT) component specific information to Wind Farm (WF) operators to be used for optimal maintenance planning. Their economic benefit to O&M costs has been investigated in [3, 4], and proven to be substantial although it largely depends on the fault detection rate [5]. While many commercial solutions, techniques and methods are available [6, 7], their related cost and complexity deter operators from a widespread deployment [8]. The use of data from the Supervisory Control And Data Acquisition (SCADA) system appears therefore as a potential solution for WT CM due to its availability at no additional cost.

The SCADA system usually samples data at relatively high frequency (typically 1 Hz) with standard practice to store 10-minute averaged values of the parameters characterising the operating and environmental conditions. The number of channels available varies considerably between manufacturers and SCADA services providers, although the minimum set typically includes wind speed and direction, active and reactive power, rotational speed, pitch and yaw angles and ambient temperature. SCADA-based monitoring may focus on individual WT components or on the whole turbine or farm. The latter may be approached through monitoring WT performance, as it is the dominant characteristic of WT operation. In general, performance monitoring is synonymous with WT Power Curve (PC) monitoring. Similar to normal behaviour models [9], operational data are used to develop a reference PC during normal operating conditions, together with related confidence intervals that allow the discrimination between normal and abnormal operation over time [10].

The analysis, modelling and monitoring of the PC has received considerable research interest over the last decade. Due to its great variability depending on site conditions [11], it is essential to accurately predict WT power output for numerous applications, such as WF energy yield prediction, wind power forecasting or monitoring and troubleshooting, as stressed in [12]. All require a

deep understanding of WT performance under real operating conditions.

The most widespread method for modelling a WT PC is the Method of bins (MOB) [13]. Although conceived for power performance testing, its use has been
55 extended and it is the current industry practice for performance monitoring and PC analysis. Nevertheless, this method relies on a linear averaging procedure leading to problems in the case of turbulent wind conditions [14]. Much effort has therefore been dedicated to develop alternative methods for WT PC modelling, mainly data-driven techniques. A comprehensive review can be found in
60 [12]. In general, PC models can be classified into parametric or non-parametric techniques. While parametric methods fit a functional form to the data by fitting one or more parameters, non-parametric models infer a functional form and do not need to make any prior assumption about the data. Some examples of parametric techniques are logistic expressions [10, 15], polynomial regressions
65 [16, 17], the Linear Hinge Model [18], or a modified hyperbolic tangent [19]. An interesting parametric method is the dynamical power curve, based on the application of the Langevin equation [14], that allows to take into account the effect of turbulent conditions. Non-parametric techniques have been significantly more investigated. The most important examples found in the literature
70 are the k -Nearest Neighbours (k -NN) algorithm [10, 20, 21], Artificial Neural Network (ANN) [20, 22, 23] and the Random Forest (RF) algorithm [24, 25, 26]. Kernel methods were also explored in [27, 28, 29]. In general, non-parametric models seem to provide higher PC modelling accuracy due to their flexibility and capability to capture features inherent in the data. Recent multivariate
75 approaches [26], including model inputs other than the wind speed, show significant accuracy improvements over univariate modelling. Nevertheless, the majority of these methods are purely deterministic, leading to a difficult assessment of their related uncertainty and therefore their related confidence intervals for normal performance characterisation. The current state-of-the-art consists
80 on the use of statistical control charts [10, 30], that are usually built based upon the modelling error metrics. Other methods applied recently, like the Gaussian Process [24, 31, 32], can provide a joint prediction of the WT power output

and its related uncertainty. Finally, copula modelling has also been explored for probabilistic monitoring of WT performance [33, 34].

85 With regards to the data used for PC modelling, most contributions rely on 10-minute averaged SCADA data. This low temporal resolution, together with the averaging effect, negatively affects the detection capabilities of some SCADA-based monitoring approaches [7]. This is of special importance for WT performance monitoring due to the rapid wind speed and power output
90 fluctuations. Indeed, the dynamic behaviour of WT performance is not yet reflected in PC modelling techniques [35] and, as a result, intermittent or transient anomalies cannot be detected. Using high-frequency SCADA data instead of averaged signals should allow dynamic turbine behaviour to be identified with higher fidelity and thus improve detection capabilities [20, 36, 37]. Although
95 some attempts have been made to investigate WT performance monitoring using SCADA data of high resolution [14, 37, 38], the subject is still at an early stage. This paper investigates in detail the potential of such high-frequency data for WT PC modelling and monitoring. The research presented here examines state-of-the-art multivariate non-parametric models and their sensitivity
100 to a range of factors, including SCADA data time resolution, for effective WT performance monitoring.

This paper is organised as follows. Section 2 details the suggested approach for WT performance monitoring, with four alternatives for PC modelling, and describes the factors to be considered in the sensitivity analysis. In Section 3,
105 the proposed methodology is applied to four different sites and detailed results of the sensitivity study are presented. Section 4 discusses the obtained results in the context of the usefulness of the developed approach for effective WT performance monitoring. Section 5 concludes the paper by highlighting the research contribution, the limitations of the presented work and discusses future
110 research needs.

2. Methodology

The methodology for WT performance monitoring is presented in Figure 1, and can be summarised in four different phases. Both operational data and alarm logs are first gathered from the SCADA system and the latter are translated into component-related information during a pre-processing phase (1). Operational data and the translated logs are then integrated and synchronised in time, to facilitate the data filtering process (2). Once the operational data have been filtered to only include normal performance representation, they are used to build a normal performance model consisting of the PC model and its related uncertainty (3). Finally, the difference between modelled and actual WT power production, the power residual, can be monitored over time (4), together with related thresholds characterising normal performance, so as to detect abnormal behaviour. These different phases of the methodology are described in greater detail below.

2.1. Alarms pre-processing

Both operational data and alarm logs from the SCADA system are utilised in the developed framework. The alarm logs usually list fault events, warnings or other relevant information. Based on a previous work published by the authors [39], a modernised WT taxonomy and original technical documentation provided by the manufacturers are used to classify the alarm logs as component-related information. This step transforms this vast volume of data into valuable information relevant to WT condition during the periods when the alarms were recorded.

2.2. Data filtering

Historical operational SCADA data used to build the normal performance modelling is filtered to ensure the representation of WTs operating under faultless conditions.

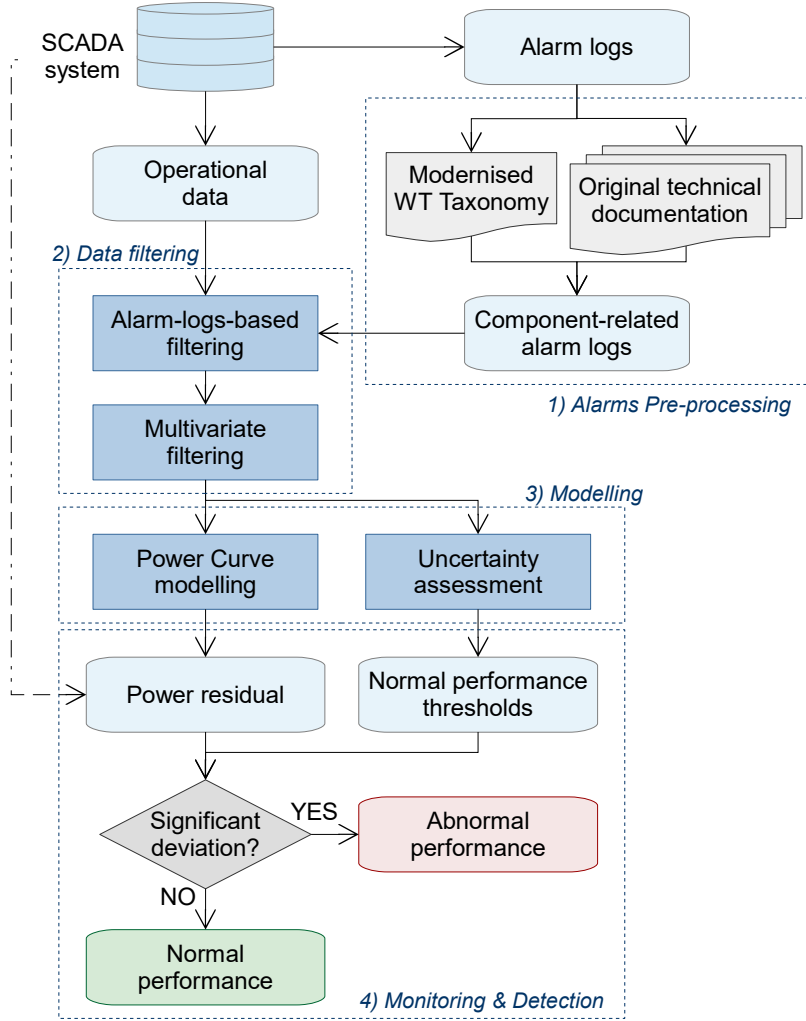


Figure 1: General framework for WT performance monitoring.

Alarm-logs-based filtering. The previously categorised alarms are integrated and synchronised in time with the operational data. In this way the operational SCADA data can be flagged for periods when the WT was experiencing an issue, and thus discarded from the training data to be used in the modelling phase. An example using high-frequency SCADA data is illustrated in Figure 2. All figures presented here are normalised for confidentiality reasons.

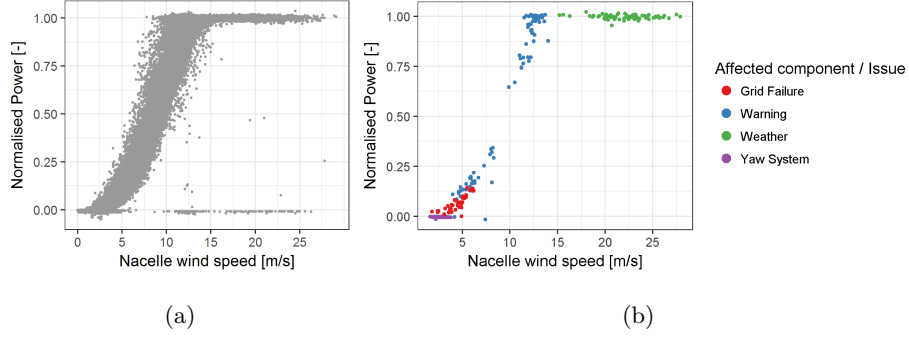


Figure 2: Example of WT normal (a) and abnormal performance (b) as indicated by the component-related SCADA alarm logs.

Multivariate filtering. Some problems might not be covered and therefore recorded
 145 by the SCADA system, as can be seen in Figure 2(a). As a result, there is a
 need for further abnormal performance filtering. Since the presented work relies
 only on the use of SCADA data, scatter PCs are built using power output and
 nacelle wind speed measurements. It is widely known that nacelle wind speed
 measurements are susceptible to high uncertainties since they are exposed to
 150 numerous sources of disturbance [40]. As a result, underperformance cannot be
 filtered by only considering nacelle wind speed data. To ensure a more robust
 filtering process, pitch angle and rotor speed measurements are also used. As
 one can see in Figure 3, the different operating regimes can be clearly iden-
 tified and so can WT normal operation. A similar approach was followed in
 155 [11], showing that incorporating the pitch angle allows flagging of data that
 could lie reasonably within the expected PC. An example of final data filtering
 is shown in Figure 4. It is important to mention that this filtering approach
 is only suitable for pitch-regulated machines, and prior knowledge about the
 control characteristics, such as the ranges of rotational speed, is key to ensuring
 160 a robust data filtering process.

2.3. Normal performance modelling

Recent research has highlighted the accuracy of multivariate non-parametric
 methods for PC modelling [26]. In line with this, three different multivariate

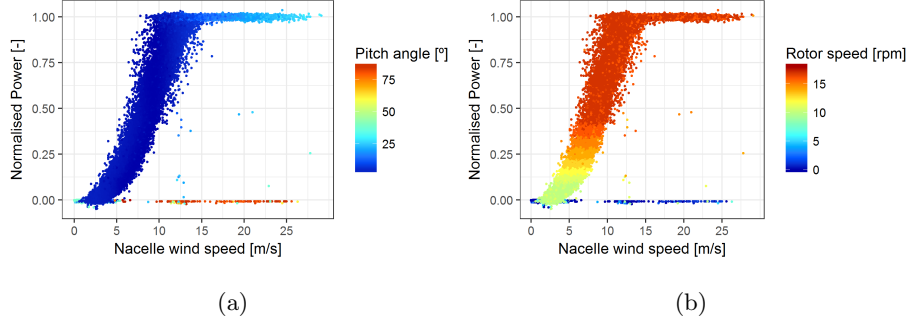


Figure 3: Example of scatter PC performance varying depending on blade pitch angle (a) and rotor speed (b).

non-parametric models, as well as the current industry practice, are applied here
 165 for modelling a reference performance or PC. Although important research has
 been conducted on the use of the ANN, this approach was discarded here due to
 its often observed overfitting problems. In this multivariate approach, nacelle
 wind speed, ambient temperature, pitch angle and rotor speed, were selected
 as input variables to predict the WT power production. All the methods are
 170 described below.

Method of bins (MOB). The MOB is the current industry practice for WT
 performance testing [13, 41]. It relies on the reduction of the data into mean
 values for individual wind speed intervals. For each of these, wind speed and
 power output data are reduced to the calculated mean, creating a univariate
 175 model that can be used for predicting the power output given a new wind speed
 observation. In this paper, data normalisation to take into account changes in
 air density is applied as detailed in [13].

k-Nearest Neighbours (kNN). The k -NN approach predicts a new sample using
 the k closest samples from the training dataset [42]. The selected metric for
 180 the distance between samples is here the Euclidean distance. The value of k
 is selected here according to the minimum Root Mean Square Error (RMSE)
 obtained from a 10-fold cross validation with the training dataset. Further
 details are provided in section 3.

Random Forests (RF). RF are considered as an ensemble method since they
 185 operate by randomly constructing multiple regression trees, avoiding the over-
 fitting problems often related to the use of a single regression tree [43]. In each
 single regression tree, the training output values are split according to conditions
 of the input values. New predictions are made by averaging the individual pre-
 dicted values from the multiple regression trees, according to conditional inputs.
 190 RF are typically tuned by selecting the number of randomly selected predictors.
 Notwithstanding the low number of predictors used here, the training dataset
 is used to tune the model following a 10-fold cross validation.

Support Vector Regression (SVR). SVR is the only kernel method considered
 here. Kernels can be understood as similarity functions [42] and in this case, a
 195 Radial Basis kernel was selected, suitable when no prior knowledge of the data
 is available. Similarly, the Radial Basis kernel parameters are selected here
 based on the minimum RMSE identified from a 10-fold cross validation with
 the training dataset.

2.3.1. Sensitivity study

200 The key contribution of this paper lies in the use of high-frequency measure-
 ments gathered from the SCADA system, rather than 10-minute averaged data.
 To investigate the potential of such high-frequency data, a sensitivity study is
 undertaken to investigate the PC modelling accuracy of the selected methods
 for different data time resolutions. Furthermore, as stressed in [27] having a
 205 better understanding of the seasonality and location effects may contribute to
 more effective WT performance monitoring. To this end, the effect of seasonal-
 ity and wind farm terrain complexity are also included as factors affecting PC
 modelling and WT monitoring capabilities. Finally, since the proposed machine
 learning algorithms for PC modelling are multivariate, the importance of input
 210 parameters is also assessed.

2.3.2. Uncertainty assessment

Since WT performance depends to a significant extent on environmental conditions, normal performance thresholds are usually defined to take account for this variability and therefore to ensure effective abnormal behaviour detection. To do so, the uncertainty related to the power predicted from the normal performance model needs to be properly addressed. In this study, two approaches are considered for assessing the uncertainty related to PC modelling.

Training error distribution metrics. As no functional form has been fitted to the data when using non-parametric models, the assessment of their related uncertainty is not straightforward. State-of-the-art methods rely on the assumption of a normal distribution of the modelling error [10, 30]. As a result, statistical control charts can be created based on the statistical metrics of the training error distribution, i.e. the mean modelling residual (μ_{train}), the standard deviation (σ_{train}) and the number of observations (N_{train}). The lower control limit (LCL) and the upper control limit (UCL) can be defined as follows, where η can be selected in order to adjust the proportion of the training residuals that are retained within the normal performance thresholds:

$$LCL = \mu_{train} - \eta \frac{\sigma_{train}}{\sqrt{N_{train}}} \quad (1)$$

$$UCL = \mu_{train} + \eta \frac{\sigma_{train}}{\sqrt{N_{train}}} \quad (2)$$

Bootstrapping. The state-of-the-art approach based on the definition of control charts uses statistical metrics from a sample dataset, that is the modelling training error. Since non-parametric regression algorithms are applied here for PC modelling, Monte Carlo technique cannot be used to produce confidence intervals for the regression errors. The bootstrapping technique is a derivation of Monte Carlo, suitable for estimates that cannot be expressed in an equation [44]. To the best of the author's knowledge, this technique has not been employed for producing the mentioned normal performance control charts, apart from few examples [24]. The produced control limits can be defined with Equation 1 and

Equation 2 as well, where μ and σ are obtained by bootstrapping instead of derived from sample statistics.

2.4. WT performance monitoring

230 As illustrated in Figure 1, the last phase consists in the identification of abnormal WT performance. The difference between the actual and the modelled power production, so-called the power residual, is monitored and compared to the normal performance thresholds. If the power residual is found to be beyond these thresholds abnormal performance is detected, covering both over
235 and underperformance situations.

3. Results

3.1. Data description

High-frequency SCADA data from four operating WFs were used to evaluate the PC modelling accuracy of the described methods. Their sensitivity
240 to site conditions, seasonality effects, input parameter variation and data time resolution was also assessed, and discussed in the context of WT performance monitoring.

Only one full year of data was available for the selected sites. As extracted from the failure documented information provided by the WF operator, only
245 healthy WTs were retained for this study. All the WTs included in this analysis are of the same model, 3-bladed turbines, gear-drive, pitch-regulated and equipped with doubly fed induction generators (DFIG).

For the sake of simplicity, the results presented in this section are illustrated through the results for a representative WT in each WF, although these were
250 consistent and of the same order of magnitude for all the WTs within each WF.

3.1.1. Wind farm terrain complexity assessment

Detailed information about the different wind farms is presented in Table 1. The effect of site conditions was addressed through the concept of terrain

complexity. Wind flow is normally distorted due to site conditions, such as topographical variations, other turbines or obstacles [41]. Since a more disturbed flow may increase the loads exerted upon the WT as well as the uncertainty in measurements and the variability of WT performance, the different WFs were classified according to the averaged terrain complexity class obtained initially per WT location. The digital height data used for this purpose was derived from the Shuttle Radar Topographic Mission (SRTM) [45], and terrain complexity was assessed following the classification defined in [41]. As one can see in Table 1, the selected sites are representative of a large range of terrain complexities from very simple to very complex.

Table 1: Wind farm information and averaged terrain complexity class.

WF	WTs	RIX class	Slope class	Overall class	Terrain Complexity
1	25	1.00	2.00	3.00	Simple
2	10	1.40	2.00	3.50	Rather complex
3	30	2.57	2.00	4.30	Complex
4	36	3.94	2.00	5.00	Very Complex

3.1.2. High resolution SCADA data

Data of wind speed, ambient temperature, power production, pitch angle and rotor speed were utilised at 0.25Hz (4 second resolution) for this study. Nacelle wind speed, ambient temperature, pitch angle and rotor speed were selected here as model inputs.

Missing data can limit the usefulness of operational SCADA data, particularly when using models with multiple input variables; in order to overcome this problem, re-sampling was applied to incomplete records to fill in the most plausible values of absent variables. A linear approach to re-sampling was applied to obtain pitch angle and rotor speed data, whereas a cubic spline re-sampling was applied to obtain ambient temperature, only available at the typical 10-minute resolution though.

3.2. Data filtering

Since only one full year of measurements was available, the first month (January) was selected for model training purposes. After the pre-processing of the alarm logs registered by the SCADA system for the selected WTs, this one month of high-frequency data and 10-minute operational data were filtered according to the process described in Section 2.2. Statistics resulting from this process are presented in Table 2. A slightly higher percentage of final observations was retained when using high-frequency data, apart from the case of the very complex site, where the difference is much more significant. The difference of the filtering results between the high-frequency and 10-minute data is also illustrated in Figure 4.

Table 2: Data filtering statistics for high-frequency and 10-minute data at each wind farm.

Data resolution	WF	Raw obs.	Alarms flagged [%]	Underperf. flagged [%]	Final obs. [%]
4-second	1	180818	0.0741	0.2063	99.7196
	2	91328	0.3153	0.5168	99.1678
	3	182428	0.9829	0.472	98.5452
	4	74250	6.0943	15.4855	78.4202
10-minute	1	4240	0.0708	1.0142	98.9151
	2	4356	0.2984	0.5969	99.1047
	3	4348	0.851	3.1509	95.9982
	4	3549	4.5647	29.304	66.1313

3.3. Model Performance

The January subset of the data was split into training (80%) and testing (20%) sets. The training (or in-sample) dataset was first used to tune the three regression models following a typical 10-fold cross validation approach, in order to avoid overfitting. This translates into a random division of the training set into 10 groups or folds of approximately equal size, where each one at a time

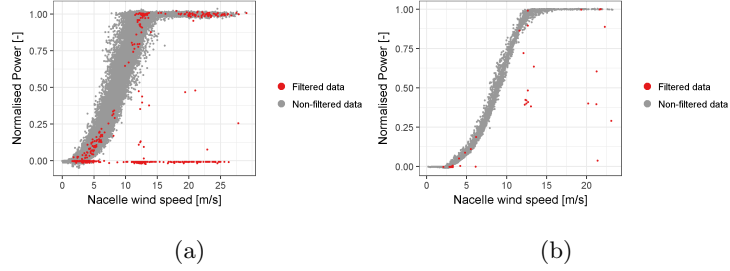


Figure 4: Example of scatter PC filtering when using high-frequency (a) and 10-minute SCADA data (b).

was treated as the validation set. The testing (or out-of-sample) dataset was used to evaluate modelling performance based on typical error metrics as the Mean Absolute Error (MAE) and the RMSE:

$$MAE = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} |\hat{y}_i - y_i| \quad (3)$$

$$RMSE = \sqrt{\frac{1}{N_{test}} \sum_{i=1}^{N_{test}} (\hat{y}_i - y_i)^2} \quad (4)$$

Model performance for each WF are presented in Table 3, expressed as a percentage of the WT rated power.

Table 3: Model accuracy results for four different WFs.

WF	MAE (%)				RMSE (%)			
	1	2	3	4	1	2	3	4
MOB	3.0568	3.0500	2.5391	5.2371	4.8975	5.5288	4.6956	8.1562
kNN	1.8267	1.9603	1.7390	4.0900	3.2596	4.0183	3.6156	7.0132
RF	1.8373	1.9287	1.7543	4.0642	3.4773	4.0430	3.7100	6.9807
SVR	2.1037	2.3542	2.1941	4.3574	3.5034	4.1841	3.8236	7.0432

290 Table 3 shows that all the multivariate non-parametric models are better at predicting WT power than the MOB. Both the k -NN and RF algorithms provide very good results, and of the same order of magnitude. This higher

accuracy is observed through both error metrics leading to not only a lower average error (related to a lower MAE), but also to a lower variance associated with the frequency distribution of error magnitudes for the multivariate non-parametric models (related to a lower RMSE). This means that they are more capable of expressing WT performance making them better predictors. This is also illustrated in Figure 5, where modelled PCs are presented for WF2.

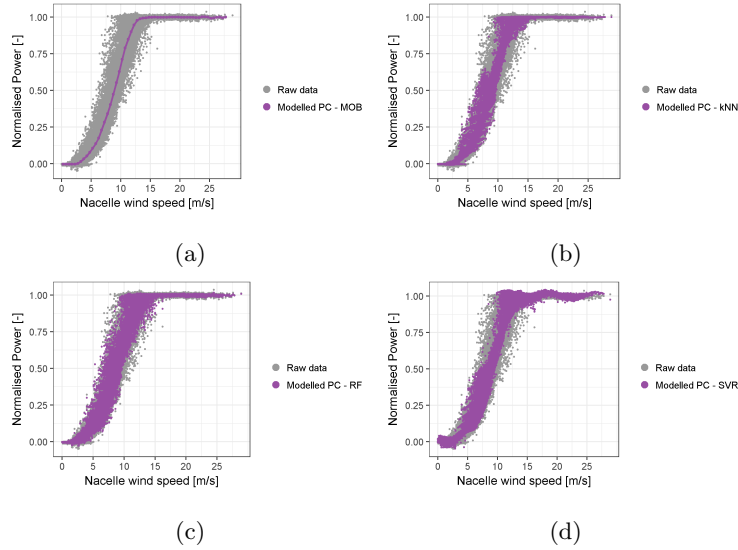


Figure 5: WT PC modelled by the MOB (a), k -NN (b), RF (c) and SVR (d).

3.4. Results from the sensitivity study

3.4.1. Wind farm terrain complexity

Results in Table 3 demonstrated that model accuracy is dependent on site. While no significant variation is observed for simple or slightly complex sites, there is an important decrease of the modelling accuracy for the very complex site, independently from the model selected. This results confirm the important dependence of modelling accuracy on site conditions.

3.4.2. Seasonality

To assess the seasonality effects, the four trained models using the high-frequency data were used to predict the rest of the year. In order not to bias results with the potential presence of abnormal performance events in other months, the same filtering procedure was applied before the WT power prediction. In this way, only WT normal performance was predicted each month. Results for the four WFs and different modelling techniques are illustrated in Figure 6 and Figure 7.

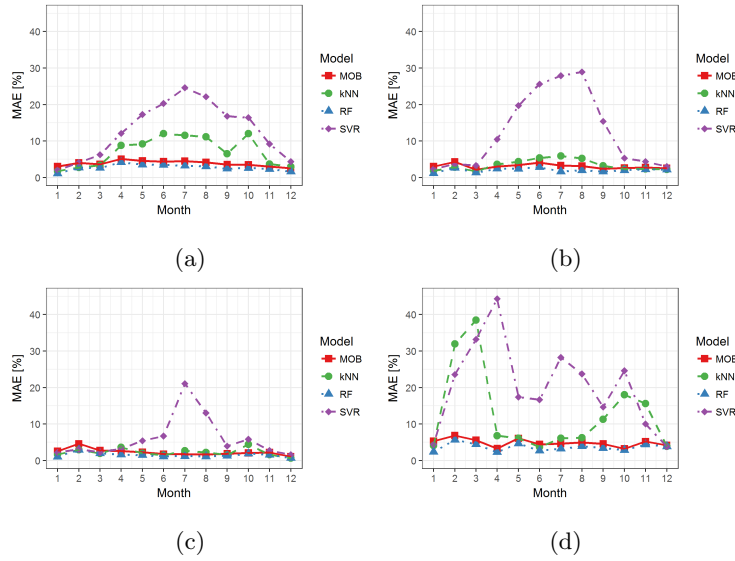


Figure 6: Seasonal variation of the MAE for the different modelling techniques at WF1 (a), WF2 (b), WF3 (c) and WF4 (d).

As can be seen, the accuracy of the MOB including the air density correction does not show any significant variation over the year, suggesting that the latter takes into account the seasonal variations of the ambient temperature and hence the air density. SVR perform significantly worse throughout the year for all the WFs considered. Even though this method includes the ambient temperature as an input and seemed to capture WT performance variability, its accuracy is hugely affected by the seasonality effects. This might be due to the choice of the

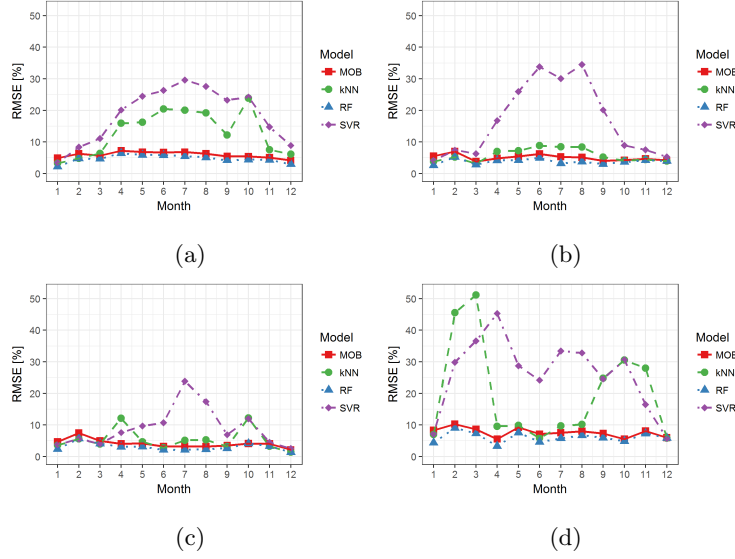


Figure 7: Seasonal variation of the RMSE for the different modelling techniques at WF1 (a), WF2 (b), WF3 (c) and WF4 (d).

kernel, which is one the greatest limitations of SVR [46]. This also applies, but less so, to the k -NN method, the variability of which also depends on the WF considered. In contrast, the RF method does not seem to be affected by any seasonality effects and preserves a very high accuracy irrespective of the month of the year.

3.4.3. Relevance of Input Variables

The use of tree-based methods, such as the RF algorithm, allows straightforward assessment of input parameter importance for multivariate modelling [26]. In this case, the relevance of the input variables was assessed through the out-of-bag (OOB) permutation importance, that measures how influential the model predictor variables are at predicting the response. The influence of a predictor increases with the value of this measure. The OOB error, evaluated here through the Mean Square Error (MSE), is the mean prediction error on each training set sample t_i using only trees that did not have t_i in their bootstrap sample. For each variable, the OOB error is estimated before and after random

permutation of the observations of the concerned variable. Then, the differences between these are averaged over all trees, and normalised by the standard deviation of the differences. If a variable highly influences predictions, then permuting its values should affect the model error. Otherwise, permutation should have little to no effect on the error. As a result, the variable importance is here
 340 quantified by the percentage rate for the increase of the MSE for each variable. Results for all the WFs considered in this study are shown in Figure 8.

While the importance of wind speed is consistent across the different sites, the importance of ambient temperature depends strongly on the WF. It also
 345 exhibits a high impact, significantly larger than wind speed. It is true that the higher the increase in MSE, the more important the variable. However, this does not imply that ambient temperature drives the greatest changes in WT power production, as this is obviously mainly driven by wind speed. It actually means that the RF model will perform significantly worse if ambient temperature is
 350 not included in the model. These results concur with the independence of RF prediction accuracy from seasonality effects, mainly due to changes in ambient temperature and air density.

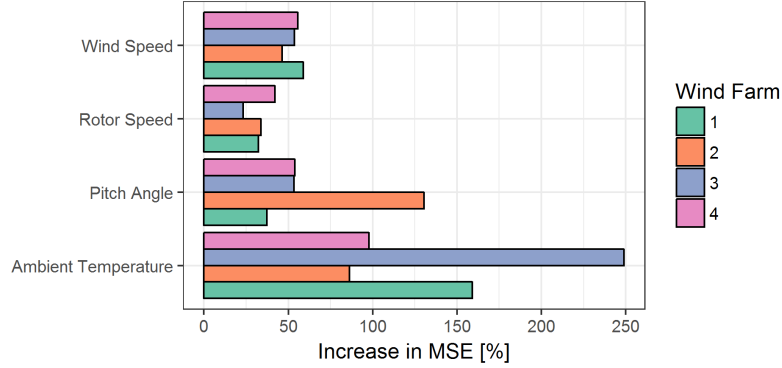


Figure 8: Site specific model input relevance.

3.4.4. Data sampling rate

One of the most important aspects of the data used in this study is its
 355 high time resolution. Indeed, this paper deals with real measurements from the

SCADA system rather than aggregated values over a 10-minute period, as is the current industry practice. To investigate the potential of using this high-frequency SCADA data two analyses were conducted. First, the autocorrelation of each parameter considered in the PC modelling was analysed individually. Then, different training datasets with corresponding different time resolutions were used to predict WT power as registered in the 10-minute aggregated data.

Data autocorrelation. When using high-frequency SCADA data one can evidently think that these data will be more affected by noise, usually smoothed by the 10-minute averaging process conventionally applied. However, this will also result in a loss of information. The averaging effect actually removes all the dynamic behaviour of WT operation. There is therefore a clear need to understand the rate of variation of the different signals considered. To this end, the autocorrelation of the nacelle wind speed, the pitch angle, the rotor speed and the power output has been calculated. The ambient temperature was omitted due to its low rate of change. The computed autocorrelations together with fitted exponential trends are shown in Figure 9. For the sake of simplicity, only results for WF2 are illustrated although consistent conclusions apply to the different sites. The trends are illustrated together in Figure 10 for comparison purposes.

WT power output shows the fastest decrease in autocorrelation, tending to zero at time lags close to the data resolution used in this study (4-second). This translates into a significantly low similarity between consecutive observations at low time resolutions, and therefore into a large loss of information regarding WT performance when aggregating power data over 10-minute periods. Wind speed and pitch angle signals show very low autocorrelation from lags of 30 seconds, although it still remains below 0.5 between lags of 4 and 30 seconds. Similarly, aggregating wind speed and pitch angle signals over 10-minute periods leads to a loss of information, but to a lesser extent than in the case of power output data. The rotor speed shows the greater autocorrelation, although it is also below 0.5 from lags of 30 seconds. This appears reasonable due to the inertia of the

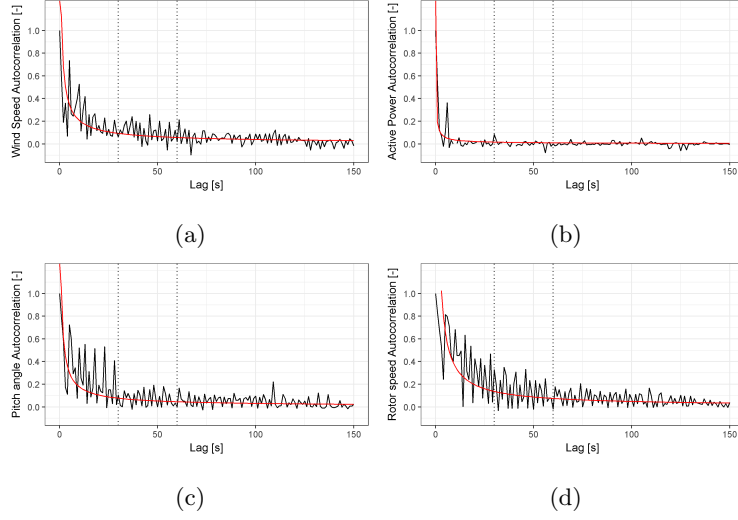


Figure 9: Autocorrelation and fitted trends for the nacelle wind speed (a), active power (b), pitch angle (c) and rotor speed (d) signals. Dotted lines indicate lags of 30 and 60 seconds respectively.

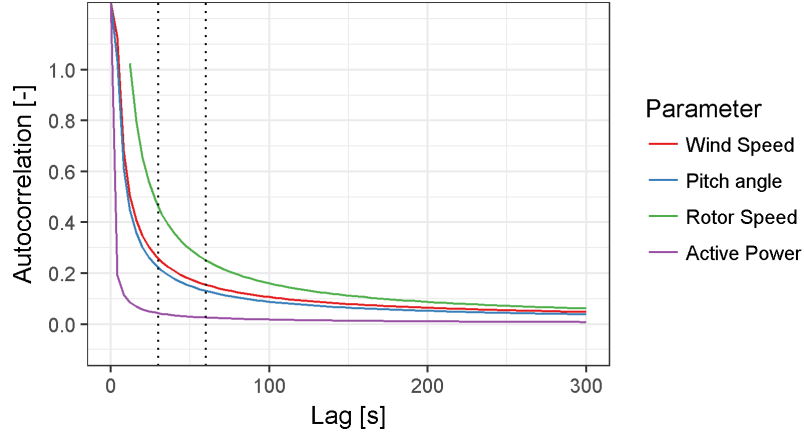


Figure 10: Autocorrelation trends for the nacelle wind speed, active power, pitch angle and rotor speed signals. Dotted lines indicate lags of 30 and 60 seconds respectively.

rotor. As a result, when considering the four signals at the same time, a time resolution of 30 seconds seems to provide a reasonable balance between time resolution and capturing WT dynamic response. If the rotor speed is omitted,

this time resolution can be further lowered.

390 *WT power prediction from different data sampling rates.* To evaluate the PC
model accuracy when using high-frequency SCADA data, different training
datasets with corresponding different time resolutions were used to train the
selected methods and to predict the power output as observed in the 10-minute
aggregated data. In this case, six different training datasets were used to pre-
395 dict the same 10-minute testing dataset; the first was the raw high-frequency
SCADA data, while the following were built based on data aggregated over pe-
riods of 30, 60, 120, 300 and 600 seconds. In a similar way to Section 3.4.2, the
power data to be predicted was filtered to solely evaluate the accuracy during
normal operating conditions. Results are illustrated in Figure 11 and Figure 12.
400 As can be seen, the MAE shows the lowest values for the 10-minute training
due to the averaging effect. However, a common trend can also be observed in
all the sites, towards lower accuracy with larger aggregated periods, with the
RF being the most significant. For all the WFs, the RMSE increases with the
aggregation period of the data, although some exceptions can be observed for
405 the 10-minute training (Figure 12 (d)). This means that, even when the purpose
is predicting at a 10-minute resolution, the highest accuracy is achieved with
highest sampling rate of the training dataset. The lower the time resolution,
the higher the variance of the error. The use of high-frequency data benefits
the understanding of WT performance variability.

410 3.5. Uncertainty assessment

Following the approach described in Section 2.3.2, the training error distri-
bution was analysed for each of the methods used for the purpose of assessing
their related uncertainty and hence building normal performance thresholds.
The combined training error distribution as well as the varying distribution
415 across different wind speed regimes are presented in Figure 13 for the four dif-
ferent models. For the sake of simplicity, these results are only given for one of
the WFs although similar results are observed for the other sites. The calculated

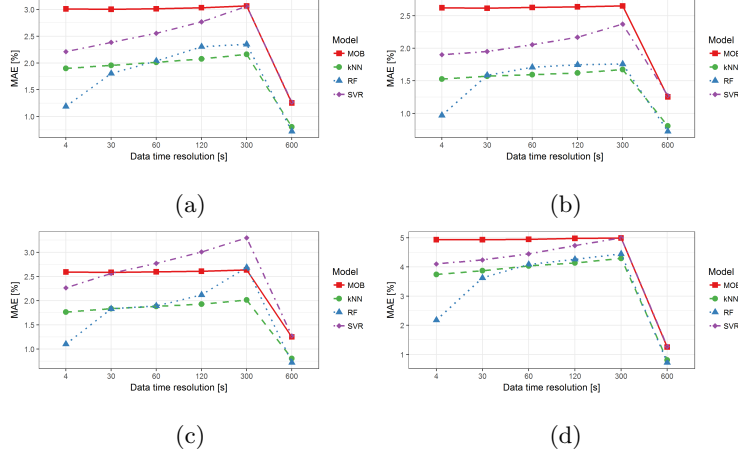


Figure 11: Prediction accuracy (MAE) for training datasets with different data time resolutions for WF1 (a), WF2 (b), WF3 (c) and WF4 (d).

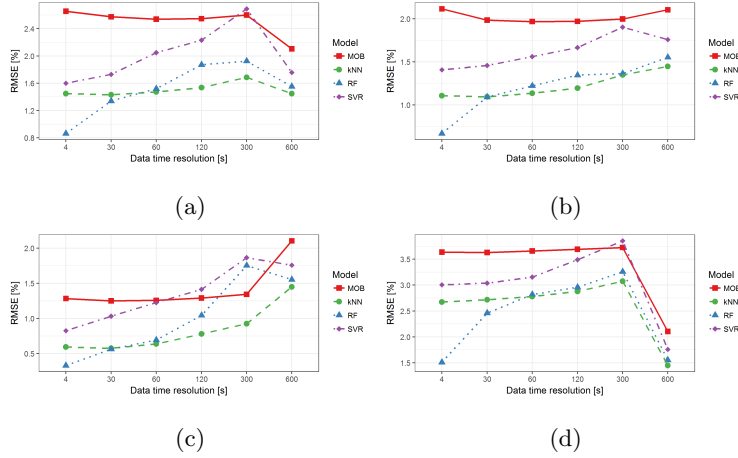


Figure 12: Prediction accuracy (RMSE) for training datasets with different data time resolutions for WF1 (a), WF2 (b), WF3 (c) and WF4 (d).

control limits from both approaches, the training error metrics and bootstrapping, were found to be very similar, although the bootstrapping derived limits were found to be slightly narrower for the same level of confidence.

Although a similar normal behaviour is observed in Figure 13 for the combined distribution, in line with the assumption made in the state-of-the-art

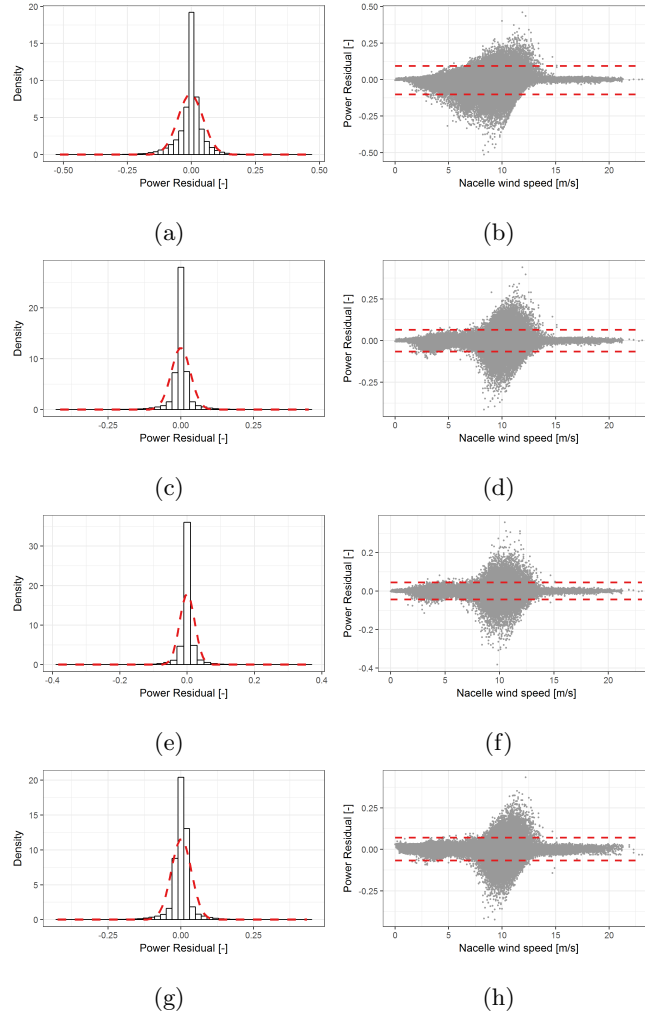


Figure 13: Training error combined distribution (left side) and varying distribution across different wind speeds (right side) for the MOB ((a),(b)), the k -NN model ((c),(d)), the RF model ((e),(f)) and the SVR model ((g),(h)). The control limits obtained from bootstrapping are illustrated with red dotted lines.

approach, an uneven envelope of the residuals across wind speed values is observed revealing a high level of heteroskedasticity in the error, as already identified [37, 38]. As the Breusch-Pagan test [47] is only suitable for linear regression models, the White test of heteroskedasticity [48] was used instead. This selected

test was run for all the modelling errors, and for the four WFs, giving a null p-value in all cases; the hypothesis of a constant variance of the error was therefore rejected in all cases under any confidence level, confirming the graphical
430 detection of the heteroskedasticity. This reveals that the assumption of a normal distribution of the error [10, 29] is invalid independently from the method selected. No matter the PC modelling accuracy, this will lead to a loss of detection capabilities during the monitoring phase. To achieve an effective monitoring of WT performance, variable performance thresholds should be produced instead.

435 4. Discussion

This section discusses the results previously presented in the context of their usefulness for effectively monitoring WT performance.

Data filtering processes have received limited attention to date, but it is crucial to ensure normal performance representation. 20 to 40 times fewer ob-
440 servations are present with the low resolution 10-minute SCADA data. This means that less data is available for training the normal performance models considering the same training period, but also that a higher percentage of data can be retained as the use of 10-minute data usually leads to over-filtered statistics. Moreover, the suggested filtering method shows robustness due to
445 the two steps involved, and to the multivariate approach. The use of high-frequency data would intuitively be expected to produce lower PC modelling accuracy, given the more complex dynamic behaviour evident at smaller time-scales and smoothed at larger granularities; however, these have been shown to be captured, learned and reproduced by multivariate non-parametric regression
450 models, especially the RF, which produces the lowest RMSE. Better detection capabilities should result, since control charts built based on lower variances should be more sensitive to deviations. Although multivariate non-parametric models, especially the k -NN and the RF, perform well and can capture WT performance variability, this ability decreases in the case of very complex flow
455 exhibited in complex sites.

The MOB with air density correction and the RF model are not affected by seasonality, unlike k -NN and SVR. SVR has greater variability, while k -NN performance depends on the site meaning both the models are unsuited to monitoring WT performance over time. The RF algorithm relied only on a short training period and accuracy remained independent from seasonality effects, making it the most suitable of the models studied here. Ambient temperature appears to play an important role in seasonal dependent modelling accuracy, as the variable importance study showed. It is therefore crucial to include it as an input variable for PC modelling to ensure accurate WT power prediction and effective performance monitoring. Seasonality effects observed for k -NN and SVR models could be avoided by taking longer training periods covering different seasons, but this will restrict the monitoring opportunities over the period selected for training and consequently delay operational application.

WT power output shows a very low autocorrelation that translates into a large loss of information when the data is aggregated over a period of 10-minutes, leading to loss of capability for detecting WT power deviations. Evidence has also been provided in this paper for higher accuracy, in terms of error variability (RMSE), for higher sampling rates of the training dataset in PC modelling. Since some WF operators might be deterred from storing high-frequency SCADA data due to larger amounts of storage space needed, it has been shown that data aggregations of 30 periods may provide a good representation of WT operation. Nonetheless, high-frequency measurements are evidently optimal for effective WT performance monitoring.

Lastly, the model uncertainty assessment has revealed high heteroskedasticity of training error in all cases meaning that PC modelling accuracy actually depends on the operating regime. Assumptions of a normally distributed model error have proven to be invalid and will lead to a loss of detection capabilities. Regardless of the method selected for normal performance modelling, control charts or anomalous behaviour thresholds should vary across the different wind speed regimes.

5. Conclusions

In this paper, the use of high-frequency SCADA data for WT performance monitoring has been thoroughly investigated. A novel framework has been proposed, based on the use of multivariate non-parametric models. A detailed
490 sensitivity study has also been conducted with real data from four different operating wind farms, to evaluate the implications and limitations derived from the use of high-frequency data. The results, discussed in a practical context, demonstrate that using high-frequency data is beneficial for performance monitoring purposes whereas the assessment of the model related uncertainty remains
495 the strongest drawback.

It has been shown that the dynamic WT behaviour, usually smoothed in the 10-minute aggregated data, is understood, learned and reproduced by multivariate non-parametric regression models, especially the RF algorithm. While other models were found to be sensitive to seasonality effects, the RF method
500 provides good accuracy on a short training period. The combination of this method with the use of high-frequency data produces the lowest model error variation. This leads to better detection capabilities, since control charts built based on lower variances will be more sensitive to deviations. Although the model accuracy depends on the site specific conditions, assessed through terrain complexity classes, the methodology is shown to be appropriate for any site
505 and location.

Nevertheless, the current state-of-the-art for the assessment of the related uncertainty for non-parametric models used in PC modelling shows the greatest drawback towards achieving an effective monitoring of WT performance. This
510 high heteroskedasticity of the training error makes the use of statistical control charts unsuitable for characterising WT normal performance. The move from purely deterministic models, as the RF algorithm, towards probabilistic methods should allow to effectively monitor WT performance regardless the mode of operation. Evidence on this issue has already been provided by the authors in
515 [38], with an empirical probabilistic assessment of WT performance.

Future work should therefore focus on the use of probabilistic methods for PC modelling in order to develop improved techniques for WT performance monitoring. Testing and evaluation of the methodology for healthy and faulty turbines are also needed to confirm its detection capabilities. Several component failures should be considered, in order to evaluate the relation between WT abnormal performance and faulty component behaviour.

Acknowledgements

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 642108. The authors gratefully acknowledge Enel Green Power for their cooperation and supply of the datasets.

References

- [1] Wind Europe, Wind energy in europe: Outlook to 2020, Available at <https://windeurope.org/about-wind/reports/wind-energy-in-europe-outlook-to-2020/> (2017/11/01).
- [2] J. Carroll, A. McDonald, I. Dinwoodie, D. McMillan, M. Revie, I. Lazakis, Availability, operation and maintenance costs of offshore wind turbines with different drive train configurations, Wind Energy 20 (2) (2017) 361–378. doi:10.1002/we.2011.
- [3] D. McMillan, G. W. Ault, Quantification of Condition Monitoring Benefit for Offshore Wind Turbines, Wind Engineering 31 (4) (2007) 267–285. doi:10.1260/030952407783123060.
- [4] J. Nilsson, L. Bertling, Maintenance Management of Wind Power Systems Using Condition Monitoring SystemsLife Cycle Cost Analysis for Two Case Studies, IEEE Transactions on Energy Conversion 22 (1) (2007) 223–229. doi:10.1109/TEC.2006.889623.

- [5] A. May, D. McMillan, S. Thöns, Economic analysis of condition monitoring systems for offshore wind turbine sub-systems, *IET Renewable Power Generation* 9 (8) (2015) 900–907. doi:10.1049/iet-rpg.2015.0019.
- 545 [6] F. P. García Márquez, A. M. Tobias, J. M. Pinar Pérez, M. Papaelias, Condition monitoring of wind turbines: Techniques and methods, *Renewable Energy* 46 (2012) 169–178. doi:10.1016/j.renene.2012.03.003.
- [7] W. Yang, P. J. Tavner, C. J. Crabtree, Y. Feng, Y. Qiu, Wind turbine condition monitoring: technical and commercial challenges, *Wind Energy* 550 17 (5) (2014) 673–693. doi:10.1002/we.1508.
- [8] J. Tautz-Weinert, S. J. Watson, Using SCADA data for wind turbine condition monitoring a review, *IET Renewable Power Generation* 11 (4) (2017) 382–394. doi:10.1049/iet-rpg.2016.0248.
- [9] M. Schlechtingen, I. Ferreira Santos, Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection, *Mechanical Systems and Signal Processing* 555 25 (5) (2011) 1849–1875. doi:10.1016/j.ymssp.2010.12.007.
- [10] A. Kusiak, H. Zheng, Z. Song, On-line monitoring of power curves, *Renewable Energy* 34 (6) (2009) 1487–1493. doi:10.1016/j.renene.2008.10.022.
- 560 [11] C. M. St. Martin, J. K. Lundquist, A. Clifton, G. S. Poulos, S. J. Schreck, Wind turbine power production and annual energy production depend on atmospheric stability and turbulence, *Wind Energy Science* 1 (2) (2016) 221–236. doi:10.5194/wes-1-221-2016.
- 565 [12] M. Lydia, S. S. Kumar, a. I. Selvakumar, G. E. Prem Kumar, A comprehensive review on wind turbine power curve modeling techniques, *Renewable and Sustainable Energy Reviews* 30 (2014) 452–460. doi:10.1016/j.rser.2013.10.030.

- [13] IEC 61400-12-1:2017, Wind energy generation systems - Part 12-1: Power
570 performance measurements of electricity producing wind turbines, Standard, International Electrotechnical Commission (2017).
- [14] T. A. Mücke, M. Wächter, P. Milan, J. Peinke, Langevin power curve
analysis for numerical wind energy converter models with new insights on
high frequency power performance, *Wind Energy* 18 (11) (2015) 1953–1971.
575 doi:10.1002/we.1799.
- [15] M. Lydia, A. I. Selvakumar, S. S. Kumar, G. E. P. Kumar, Advanced
Algorithms for Wind Turbine Power Curve Modeling, *IEEE Transactions
on Sustainable Energy* 4 (3) (2013) 827–835. doi:10.1109/TSTE.2013.
2247641.
- 580 [16] O. Uluyol, G. Parthasarathy, W. Foslien, K. Kim, Power Curve Analytic
for Wind Turbine Performance Monitoring and Prognostics, in: Annual
Conference of the Prognostics and Health Management Society, 2011.
- [17] S. Shokrzadeh, M. Jafari Jozani, E. Bibeau, Wind Turbine Power Curve
Modeling Using Advanced Parametric and Nonparametric Methods, *IEEE
585 Transactions on Sustainable Energy* 5 (4) (2014) 1262–1269. doi:10.1109/
TSTE.2014.2345059.
- [18] R. J. de Anadrade Vieira, M. A. Sanz-bobi, Power Curve Modelling of a
Wind Turbine for monitoring its behaviour, in: 4th International Confer-
ence on Renewable Energy Research and Applications, 2015.
- 590 [19] E. Taslimi-Renani, M. Modiri-Delshad, M. F. M. Elias, N. A. Rahim, De-
velopment of an enhanced parametric model for wind turbine power curve,
Applied Energy 177 (2016) 544–552. doi:10.1016/j.apenergy.2016.05.
124.
- [20] M. Schlechtingen, I. F. Santos, S. Achiche, Using Data-Mining Approaches
595 for Wind Turbine Power Curve Monitoring: A Comparative Study, *IEEE*

Transactions on Sustainable Energy 4 (3) (2013) 671–679. doi:10.1109/TSTE.2013.2241797.

- [21] G. A. Skrimpas, K. Kleani, N. Mijatovic, C. W. Sweeney, B. B. Jensen, J. Holboell, Detection of icing on wind turbine blades by means of vibration and power curve analysis, Wind Energy 19 (10) (2016) 1819–1832. doi:10.1002/we.1952.
- [22] A. Marvuglia, A. Messineo, Monitoring of wind farms’ power curves using machine learning techniques, Applied Energy 98 (2012) 574–583. doi:10.1016/j.apenergy.2012.04.037.
- [23] F. Pelletier, C. Masson, A. Tahan, Wind turbine power curve modelling using artificial neural network, Renewable Energy 89 (2016) 207–214. doi:10.1016/j.renene.2015.11.065.
- [24] V. Bulaevskaya, S. Wharton, a. Clifton, G. Qualley, W. O. Miller, Wind power curve modeling in complex terrain using statistical models, Journal of Renewable and Sustainable Energy 7 (1) (2015) 013103. doi:10.1063/1.4904430.
- [25] A. Clifton, L. Kilcher, J. K. Lundquist, P. Fleming, Using machine learning to predict wind turbine power output, Environmental Research Letters 8 (2) (2013) 024009. doi:10.1088/1748-9326/8/2/024009.
- [26] O. Janssens, N. Noppe, C. Devriendt, R. V. de Walle, S. V. Hoecke, Data-driven multivariate power curve modeling of offshore wind turbines, Engineering Applications of Artificial Intelligence 55 (2016) 331–338. doi:10.1016/j.engappai.2016.08.003.
- [27] G. A. Skrimpas, C. W. Sweeney, K. S. Marhadi, B. B. Jensen, N. Mijatovic, J. Holboll, Employment of Kernel Methods on Wind Turbine Power Performance Assessment, IEEE Transactions on Sustainable Energy 6 (3) (2015) 698–706. doi:10.1109/TSTE.2015.2405971.

- [28] N. Yampikulsakul, E. Byon, S. Huang, S. Sheng, M. You, Condition Monitoring of Wind Power System With Nonparametric Regression Analysis, *IEEE Transactions on Energy Conversion* 29 (2) (2014) 288–299. doi:10.1109/TEC.2013.2295301.
- [29] T. Ouyang, A. Kusiak, Y. He, Modeling wind-turbine power curve: A data partitioning and mining approach, *Renewable Energy* 102 (2017) 1–8. doi:10.1016/j.renene.2016.10.032.
- [30] P. Cambon, R. Lepvrier, C. Masson, a. Tahan, F. Pelletier, Power curve monitoring using weighted moving average control charts, *Renewable Energy* 94 (2016) 126–135. doi:10.1016/j.renene.2016.03.031.
- [31] E. Papatheou, N. Dervilis, A. E. Maguire, C. Campos, I. Antoniadou, K. Worden, Performance monitoring of a wind turbine using extreme function theory, *Renewable Energy* 113 (2017) 1490–1502. doi:10.1016/j.renene.2017.07.013.
- [32] Jin Zhou, Peng Guo, Xue-Ru Wang, Modeling of wind turbine power curve based on Gaussian process, in: 2014 International Conference on Machine Learning and Cybernetics, IEEE, 2014, pp. 71–76. doi:10.1109/ICMLC.2014.7009094.
- [33] B. Stephen, S. J. Galloway, D. McMillan, D. C. Hill, D. G. Infield, A Copula Model of Wind Turbine Performance, *IEEE Transactions on Power Systems* 26 (2) (2011) 965–966. doi:10.1109/TPWRS.2010.2073550.
- [34] Y. Wang, D. G. Infield, B. Stephen, S. J. Galloway, Copula-based model for wind turbine power curve outlier rejection, *Wind Energy* 17 (11) (2014) 1677–1688. doi:10.1002/we.1661.
- [35] G. A. M. van Kuik, J. Peinke, R. Nijssen, D. Lekou, J. Mann, J. N. Sørensen, C. Ferreira, J. W. van Wingerden, D. Schlipf, P. Gebraad, H. Polinder, A. Abrahamsen, G. J. W. van Bussel, J. D. Sørensen, P. Tavner, C. L. Bottasso, M. Muskulus, D. Matha, H. J. Lindeboom,

- S. Degraer, O. Kramer, S. Lehnhoff, M. Sonnenschein, P. E. Sørensen, R. W. Künnike, P. E. Morthorst, K. Skytte, Long-term research challenges in wind energy – a research agenda by the European Academy of Wind Energy, *Wind Energy Science* 1 (1) (2016) 1–39. doi:10.5194/wes-1-1-2016.
- 655 [36] D. Coronado, K. Fischer, Condition Monitoring of Wind Turbines : State of the Art , User Experience and Recommendations, Tech. rep., Fraunhofer Institute for Wind Energy and Energy System Technology (January 2015).
- [37] E. Gonzalez, J. J. Melero, Wind turbine fault detection by monitoring its performance using high frequency SCADA data, in: 30th International Congress & Exhibition on Condition Monitoring and Diagnostic Engineer-
660 ing, COMADEM, 2017.
- [38] E. Gonzalez, B. Stephen, D. Infield, J. J. Melero, On the use of high-frequency SCADA data for improved wind turbine performance monitoring, *Journal of Physics: Conference Series* 926 (2017) 012009. doi:10.1088/1742-6596/926/1/012009.
665
- [39] E. Gonzalez, M. Reder, J. J. Melero, SCADA alarms processing for wind turbine component failure detection, *Journal of Physics: Conference Series* 753 (7) (2016) 72019. doi:10.1088/1742-6596/753/7/072019.
- [40] S. Frandsen, J. N. Sørensen, R. Mikkelsen, T. F. Pedersen, I. Antoniou, K. Hansen, The generics of wind turbine nacelle anemometry, in: European Wind Energy Conference and Exhibition 2009, 2009.
670
- [41] IEC 61400-12-2:2013, Wind energy generation systems - Part 12-2: Power performance of electricity-producing wind turbines based on nacelle anemometry, Standard, International Eletrotechnical Comission (2013).
- 675 [42] I. H. Witten, E. Frank, Data Mining: Practical Machine Learning Tools and Techniques, 2nd Edition, Morgan Kaufmann Publishers, San Francisco, CA, 2005.

- [43] L. Breiman, Random Forests, *Machine Learning* 45 (1) (2001) 5–32. doi:
10.1023/A:1010933404324.
- 680 [44] A. C. Davison, D. V. Hinkley, *Bootstrap methods and their appli-*
cation, Cambridge University Press, Cambridge, 1997. doi:10.1017/
CB09780511802843.
- [45] A. Koch, C. Heipke, Quality assessment of digital surface models derived
from the Shuttle Radar Topography Mission (SRTM), in: *Geoscience and*
685 *Remote Sensing Symposium*, 2001. IGARSS '01. IEEE 2001 International,
Vol. 6, IEEE, 2001, pp. 2863–2865. doi:10.1109/IGARSS.2001.978187.
- [46] C. J. C. Burges, A Tutorial on Support Vector Machines for Pattern
Recognition, *Data mining and knowledge discovery* 2 (2) (1998) 121–167.
doi:10.1023/A:1009715923555.
- 690 [47] T. S. Breusch, A. R. Pagan, A Simple Test for Heteroscedasticity and
Random Coefficient Variation, *Econometrica* 47 (5) (1979) 1287–1294. doi:
10.2307/1911963.
- [48] H. White, A Heteroskedasticity-Consistent Covariance Matrix Estimator
and a Direct Test for Heteroskedasticity, *Econometrica* 48 (4) (1980) 817–
695 838. doi:10.2307/1912934.